



Europäisches Patentamt
European Patent Office
Office européen des brevets



⑪ Publication number:

0 381 693 B1

⑫

EUROPEAN PATENT SPECIFICATION

⑮ Date of publication of patent specification: 29.11.95 ⑯ Int. Cl. C12Q 1/68

⑯ Application number: 88909162.5

⑰ Date of filing: 16.09.88

⑱ International application number:
PCT/US88/03194

⑲ International publication number:
WO 89/03432 (20.04.89 89/09)

④ METHOD FOR RAPID BASE SEQUENCING IN DNA AND RNA.

⑩ Priority: 07.10.87 US 105375

⑪ Date of publication of application:
16.08.90 Bulletin 90/33

⑫ Publication of the grant of the patent:
29.11.95 Bulletin 95/48

⑬ Designated Contracting States:
DE FR GB

⑭ References cited:
EP-A- 0 251 575
US-A- 3 730 844
US-A- 4 521 509

Nucleic Acids Research, Vol. 14, No. 6, 1986,
Harvard University, BAUMLEIN et al., pp.
2707-2720

DNA, Vol. 5, No. 3, 1986, Medical College of
Wisconsin, Milwaukee, WI, SHIMKUS et al,
'Synthesis and Characterization of Biotin-
Labeled Nucleotide Analogs', pp.247-255.

⑤ Proprietor: UNITED STATES DEPARTMENT OF
ENERGY
1000 Independence Avenue S.W.
Washington, D.C. 20585 (US)

⑥ Inventor: JETT, James, Hubert
545 Navajo Road
Los Alamos, NM 87544 (US)
Inventor: KELLER, Richard, Alan
4 La Rosa Court
Los Alamos, NM 87544 (US)
Inventor: MARTIN, John, Calvin
4987 Trinity Drive
Los Alamos, NM 87544 (US)
Inventor: MOYZIS, Robert, Keith
865 Los Pueblos
Los Alamos, NM 87544 (US)
Inventor: SHERA, Edgar, Brooks
4 Timber Ridge Road
Los Alamos, NM 87544 (US)
Inventor: STEWART, Carleton, Colburn
30 Carlton Drive
Hamburg, NY 14075 (US)

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

EP 0 381 693 B1

Nature, Vol. 321, 12 June 1986, Foster City,
Calif, SMITH et al, "Fluorescence Detection
In Automated DNA Sequence Analysis", pp.
674-679

Proc. Natl. Acad. Sci., Vol. 78, No. 11, New
Haven, Conn, LANGER et al "Enzymatic Syn-
thesis of Biotin-Labeled Polynucleotides:
Novel Nucleic Acid Affinity Probes", pp.
6633-6637, November 1981.

Inventor: RATLIFF, Robert, Lafayette
252 La Cueva
Los Alamos, NM 87544 (US)

④ Representative: Wagner, Karl H. et al
WAGNER & GEYER
Patentanwälte
Gewürzmühlestrasse 5
D-80538 München (DE)

Description**BACKGROUND OF THE INVENTION**

This invention is generally related to DNA and RNA sequencing and, more particularly, to DNA and RNA sequencing by detecting individual nucleotides.

A world-wide effort is now in progress to analyze the base sequence in the human genome. The magnitude of this task is apparent, with 3×10^9 bases in the human genome, and available base sequencing rates are about 200-500 bases per 10-24 hour period. Considerable interest also exists in nucleic acid sequencing from non-human sources. Existing procedures are labor intensive and cost approximately \$1 per base.

By way of example, Sanger et al., "DNA Sequencing with Chain-Terminating Inhibitors," *Proceedings of the National Academy of Science, USA* 74, 5463-7 (1977) provide for sequencing 15-200 nucleotides from a priming site. Radioactive phosphorus is used in the primer extension to provide a marker. Enzymatic resynthesis coupled with chain terminating precursors are used to produce DNA fragments which terminate randomly at one of the four DNA bases: adenine (A), cytosine (C), guanine (G), or thymine (T). The four sets of reaction products are separated electrophoretically in adjacent lanes of a polyacrylamide gel. The migration of the DNA fragments is visualized by the action of the radioactivity on a photographic film. Careful interpretation of the resulting band patterns is required for sequence analysis. This process typically takes 1-3 days. Further, there are problems with band pile-ups in the gel, requiring further confirmatory sequencing.

In a related technique, A.M. Maxam and W. Gilbert, "A New Method for Sequencing DNA," *Proceedings of the National Academy of Science, USA* 74, 560-564 (1977), teach a chemical method to break the DNA into four sets of random length fragments, each with a defined termination. Analysis of the fragments proceeds by electrophoresis as described above. The results obtained using this method are essentially the same as the "Sanger Method."

In another example, Smith et al., "Fluorescent Detection in Automated DNA Sequence Analysis," *Nature* 321, 674-679 (June 1986), teach a method for partial automation of DNA sequence analysis. Four fluorescent dyes are provided to individually label DNA primers. The Sanger method is used to produce four sets of DNA fragments which terminate at one of the four DNA bases with each set characterized by one of the four dyes. The four sets of reaction products, each containing many identical DNA fragments, are mixed together and

placed on a polyacrylamide gel column. Laser excitation is then used to identify and characterize the migration bands of the labeled DNA fragments on the column where the observed spectral properties of the fluorescence are used to identify the terminal base on each fragment. Sequencing fragments of up to 400 bases has been reported. Data reliability can be a problem since it is difficult to uniquely discern the spectral identity of the fluorescent peaks.

U.S. Patent 3,730,844 teaches a method for the analysis of polynucleotide sequences upon sequential degradation. A chemical process is used to degrade large amounts (nanomoles) of polynucleotide from which nucleoside units can be identified. Only limited amounts of sequence data can be obtained, however, since each cycle requires 100% reaction and complete stoppage at each step in order to maintain register for nucleoside identification since individual fragments are not sequenced. This technique is also exceedingly slow, i.e., on the order of 2-3 bases per hour.

These and other problems in the prior art are addressed by the present invention and an improved process is provided for rapid sequencing of DNA bases. As herein described, the present invention provides for the sequential detection of individual nucleotides cleaved from a single DNA or RNA fragment.

Accordingly it is an object of the present invention to provide an automated base sequence analysis for DNA and RNA.

Another object of the present invention is to process long strands of DNA or RNA i.e., having thousands of bases.

One other object is to rapidly sequence and identify individual bases.

SUMMARY OF THE INVENTION

To achieve the foregoing and other objects and in accordance with the purpose of the present invention as embodied and broadly described herein a method for DNA and RNA base sequencing is provided. A single fragment from a strand of DNA or RNA is hybridized to a substrate having a single site effective to hybridize with the fragment of DNA or RNA and suspended in a flowing sample stream. Using an exonuclease, the end base on the DNA or RNA fragment is repetitively cleaved from the fragment to form a train of the base in the flowing sample stream. The bases are thereafter detected in sequential passage through a detector to reconstruct the base sequence of the DNA or RNA fragment.

In another characterization of the present invention, fragments of DNA or RNA are formed from the constituent bases, which have identifiable flu-

orescent characteristics. Each fragment is hybridized to a substrate having a single site effective to hybridize with the fragment. The bases are sequentially cleaved with exonuclease from the end of a single fragment of the strands to form a train of the identifiable bases in a flowing sample stream. The single, cleaved bases in the train are then sequentially identified to reconstruct the base sequence of the DNA or RNA strand.

In one particular characterization of the invention, each of the nucleotides effective for DNA and RNA resynthesis is modified to possess an identifiable characteristic. A fragment of DNA is synthesized from the modified nucleotides where the synthesized fragment is complementary to a DNA or RNA fragment having a base sequence to be determined. A single fragment of the complementary DNA or RNA is hybridized to a substrate having a single site effective to hybridize with the DNA or RNA fragment and suspended in a flowing sample stream. Individual identifiable nucleotides are sequentially cleaved from the free end of the suspended DNA fragment with exonuclease. The single bases are then sequentially identified. The base sequence of the parent DNA or RNA strand can then be determined from the complementary DNA strand base sequence.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and form a part of the specification, illustrate an embodiment of the present invention and, together with the description, serve to explain the principles of the invention. In the drawings:

FIGURE 1 is a graphic illustration of a DNA sequencing process according to the present invention.

FIGURE 2 is a graphical representation of an output signal according to the present invention.

DETAILED DESCRIPTION OF THE INVENTION

According to the present invention, a method is provided for sequencing the bases in large DNA or RNA fragments by isolating single DNA or RNA fragments in a moving stream and then individually cleaving single bases into the flow stream, forming a sequence of the bases through a detection device. In one embodiment, the single bases in the flowing sample streams are interrogated by laser-induced fluorescence to determine the presence and identity of each base.

It will be understood that DNA and RNA strands are each formed from nucleotides comprising one of four organic bases: adenine, cytosine, quanine, and thymine (DNA) or uracil (RNA). The DNA and RNA nucleotides are similar, but not

identical; however, the nucleotides and strands of nucleotides can be functionally manipulated in a substantially identical manner. Also, the complement of an RNA fragment is conventionally formed as a DNA strand with thymine in place of uracil. The following description is referenced to DNA sequencing, but any reference to DNA includes reference to both DNA and RNA and without any limitation to DNA.

10 In a particular embodiment of the present invention, the initial step is an enzymatic synthesis of a strand of DNA, complementary to a fragment to be sequenced, with each base containing a fluorescent tag characteristic of the base. Sequencing the complementary strand is equivalent to sequencing the original fragment. The synthesized strand is then suspended in a flowing sample stream containing an exonuclease to cleave bases sequentially from the free end of the suspended DNA or RNA. The cleaved, fluorescently labeled bases then pass through a focused laser beam and are individually detected and identified by laser-induced fluorescence.

15 The maximum rate that bases may be sequenced is determined by the kinetics of the exonuclease reaction with DNA or RNA and the rate of detection. A projected rate of 1000 bases/sec would result in sequencing 8×10^7 bases/day. This is in contrast to standard techniques which take 10-24 hours to sequence 200-500 bases.

20 Referring now to Figure 1, one effective sequencing method comprises the following steps: (1) prepare a selected strand of DNA 10, in which individual bases are provided with an identifiable characteristic, e.g., labeled with color-coded fluorescent tags to enable each of the four bases to be identified, (2) select and suspend 40 a single fragment of DNA with identifiable bases in a flowing sample stream, (3) sequentially cleave 20 the identifiable bases from the free end of the suspended DNA fragment, and (4) identify the individual bases in sequence, e.g., detect 34 the single, fluorescently labeled bases as they flow through a focused laser system. Exemplary embodiments of the individual process steps are hereinafter discussed.

Selection of DNA Fragment to be Sequenced

25 In accordance with the present process, a single DNA fragment 10a is selected and prepared for labeling and analysis. In an exemplary selection process from a heterogeneous mixture of DNA fragments, avidin is bound to microspheres and a biotinylated probe, complementary to some sequence within the desired DNA fragment 10a, is bound to the avidin on the microspheres. The avidin-biotinylated probe complex is then mixed

with the heterogeneous mixture of DNA fragments to hybridize with the desired fragments 10a. The beads are separated from the unbound fragments and washed to provide the desired homogeneous DNA fragments 10a.

The selected fragments are further processed by removing the first microsphere and ligating a tail of known sequence 9 to the primer 12 attached to the 3' end of the fragment 10a. Microspheres 40 are prepared with phycoerythrin-avidin and sorted to contain a single molecule of phycoerythrin-avidin. A single complementary probe 9a to the known sequence 9 is biotinylated and bound to the sorted microspheres 40. The bead-probe complex is then hybridized to the selected fragment 10a. Thus, a single fragment of DNA 10a will be bound to each microsphere.

In another embodiment, a homogeneous source of DNA fragments is provided, e.g. from a gene library. A selection step is not then required and the homogeneous DNA fragments can be hybridized with the microspheres 40 containing a single molecule of phycoerythrin-avidin, with the appropriate complementary probe attached as above.

In either case, a single microsphere 40 can now be manipulated using, for example, a microinjection pipette to transfer a single fragment strand for labeling and analysis as discussed below.

Fluorescence Labeling of Bases

The bases forming the single fragment to be analyzed are provided with identifiable characteristics. The identifiable characteristic may attach directly to each nucleotide of DNA strand 10a. Alternatively, bases may first be modified to obtain individual identifiable characteristics and resynthesized to selected strand 10a to form a complementary DNA strand. In either event, DNA fragment 10 is provided for analysis with identifiable bases.

In one embodiment, a fluorescent characteristic is provided. The bases found in DNA do have intrinsic fluorescence quantum yields $<10^{-3}$ at room temperature. In order to detect these bases by a fluorescence technique, however, it is desirable to modify them to form species with large fluorescence quantum yields and distinguishable spectral properties, i.e., to label the bases.

Methods for synthesizing a complementary strand of DNA by an enzymatic procedure using labeled nucleotides via linker arms are known in the art. See, e.g., P. R. Langer et al., "Enzymatic Synthesis of Biotin-Labeled Polynucleotides: Novel Nucleic Acid Affinity Probes," Proc. Natl. Acad. Sci. USA 78, 6633 (1981); M. L. Shimkus et al., "Synthesis and Characterization of Biotin-Labeled

Nucleotide Analogs," DNA 5, 247 (1986). Referring to Figure 1, a primer 12 is attached to the 3' end of a DNA fragment 10a and an enzyme, e.g., DNA polymerase-Klenow, fragment, is used to synthesize the complement to DNA fragment 10a starting from the end of primer 12. Modified deoxyribonucleotides 14, 16, 18, 22 are used in the synthesis (typically modified dATP 14a, dTTP (or dUTP) 16a, dCTP 18a, and dGTP 22a).

Each of the modified nucleotides is formed with a long carbon chain linker arm 14b, 16b, 18b, and 22b, respectively, terminating in a characteristic fluorescent dye 14c, 16c, 18c, and 22c. The modified nucleotides 14, 16, 18, and 22 are then incorporated into the synthesized fragment by DNA polymerase. The long linker arms 14b, 16b, 18b, 22b isolate the fluorescent dye tags 14c, 16c, 18c, 22c from the bases 14a, 16a, 18a, 22a to permit uninhibited enzyme activity.

DNA fragments several kB long have been synthesized with each base containing a carbon chain linker arm terminating in biotin as hereinafter described. To exemplify the DNA synthesis, tagging, and cleaving processes a known strand of DNA nucleotides was formed. Nucleotides were tagged with a linker arm terminating in biotin, and a complementary strand of DNA was synthesized from the tagged nucleotides. Biotin was used as a model tag rather than fluorescent dyes to demonstrate the synthesis and cleavage reactions.

1. Preparation of known strand [d(A,G)]:

A polydeoxyribonucleotide, d(A,G)₂₁₃₈, was prepared by the method outlined in R. L. Ratliff et al., "Heteropolynucleotide Synthesis with Terminal Deoxyribonucleotidyltransferase," Biochemistry 6, 851 (1967) and "Heteropolynucleotides Synthesized with Terminal Deoxyribonucleotidyltransferase. II. Nearest Neighbor Frequencies and Extent of Digestion by Micrococcal Deoxyribonuclease," Biochemistry 7, 412 (1968). The subscript 2138, refers to the average number of bases in the fragment and the comma between the A and the G indicates that the bases are incorporated in a random order.

Ten micromoles of the 5'-triphosphate of 2'-deoxyadenosine (dATP) were mixed with one micromole of the 5'-triphosphate of 2'-deoxyguanosine (dGTP) and 5.5 nanomoles of the linear heptamer of 5'-thymidylic acid (d(pT)₇) which acts as a primer. Ten thousand units of terminal transferase were added to the solution which was buffered at pH 7 and the reaction mixture was maintained at 37°C for 24 hours. (One unit is defined as the amount of enzyme which will polymerize 1 nanomole of nucleotide in one hour.) The resulting d(A,G)₂₁₃₈ was then separated from the reaction

mixture and purified.

2. Preparation of biotinylated complementary strand $d(C,U)_{2138}$:

The complementary strand of DNA to $d(A,G)_{2138}$, prepared as described above, was synthesized from nucleotides (dCTP) and d(UTP) tagged with biotin. A mixture of 10 nanomoles of the biotinylated 5'-triphosphate of 2'-deoxycytidine (dCTP) and 20 nanomoles of the biotinylated 5'-triphosphate of 2'-deoxyuridine (dUTP) was added to 10 nanomoles of $d(A,G)_{2138}$ and 22 picomoles of d(pT)₇. Ten units of DNA polymerase (E coli), Klenow fragment, were then added to the mixture which was buffered at pH 8 and maintained at a temperature of 37°C for 2 hours. Analysis of the resulting products by electrophoresis demonstrated that the reaction went to completion and the completely biotinylated complementary DNA fragment, $d(C,U)_{2138}$, was formed.

3. Exonuclease cleavage of biotinylated $d(C,U)_{2138}$:

The completely biotinylated $d(C,U)_{2138}$, synthesized as described above, was sequentially cleaved by adding 10 units of exonuclease III to 5 nanomoles of $d(A,G)_{2138}$ + biotinylated $d(C,U)_{2138}$. The reaction mixture was maintained at pH 8 and 37°C for two hours. At the end of two hours, analysis of the reaction mixture showed that 30% of the DNA was cleaved and the cleavage reaction appeared to be still proceeding. A control reaction using normal $d(C,T)_{2138}$ yielded 85% cleavage in two hours. Hence, biotinylation does appear to slow the cleavage reaction using exonuclease III, but the tagged nucleotides were sequentially cleaved from the DNA fragments.

In accordance with the present invention, the selected fluorescent dyes are substituted for biotin to specifically tag each nucleotide type with a dye characteristic of that nucleotide. The resulting complementary DNA chain will then provide each base with a characteristic, strongly fluorescing dye. By way of example, Smith et al., *supra*, teach a set of four individually distinguishable tags.

The sensitivity for fluorescence detection can be increased, if necessary, by attaching several dye molecules along the linker arm. Alternatively, large phycoerythrin-like molecules or even small microspheres containing many dye molecules may be attached to the linker arm. In yet another alternative, fluorescent labels might be attached to the primary, single stranded fragment, thereby eliminating the necessity of forming labeled bases and synthesizing the complementary strand.

It should be noted that DNA fragment 10 may be either a single or double strand of DNA. A

single strand of DNA arises where the selected DNA strand is directly tagged for base identification or where the resynthesized complementary tagged DNA strand is separated from the selected strand. A double strand arises where the resynthesized DNA strand remains combined with the selected strand. As used herein, the term "fragment" refers to any and all of such conditions.

10 Enzymatic Cleavage of the Tagged Nucleotides

After DNA fragment 10 is formed with identifiable bases and hybridized to microsphere 40, a single fragment 10 can be manipulated with microsphere 40 and suspended in flow stream 24. Exonuclease 20 is used to cleave bases 14a, 16a, 18a, 22a sequentially from single DNA fragment 10 suspended in flow stream 24. While the presence of the linker arm and the fluorescent dye may inhibit the enzymatic activity of some exonucleases, suitable exonucleases will cleave with only a slight reduction in rate. Individual bases have been sequentially enzymatically cleaved from DNA fragments formed completely from biotinylated nucleotides as demonstrated above. See, also, e.g., M. L. Shimkus et al., *supra*. The rate of cleavage can be adjusted by varying the exonuclease concentration, temperature, or by the use of poisoning agents. The time to remove one base can be made to be on the order of one millisecond. See, e.g., W. E. Razzell et al., "Studies on Polynucleotides," *J. Bio. Chem.* 234 No 8, 2105-2112 (1959).

35 Single Molecule Detection

The individual modified nucleotides 14, 16, 18, and 22 are carried by flow stream 24 into flow cell 26 for detection and analysis by single molecule detection system 34. One embodiment of a laser-induced fluorescence detection system is described in D. C. Nguyen et al., "Ultrasensitive Laser-Induced Fluorescence Detection in Hydrodynamically Focused Flows," *J. Opt. Soc. Am. B* 4, 138-143, No. 2 (1987). The photomultiplier-based detection system described therein has detected single molecules of phycoerythrin in focused, flowing sample streams by laser-induced fluorescence. See D. C. Nguyen et al., "Detection of Single Molecules of Phycoerythrin in Hydrodynamically Focused Flows by Laser-Induced fluorescence," *Anal. Chem.* 59, 2158-2161 (September 1987).

Phycoerythrin is a large protein containing the equivalent of 25 rhodamine-6G dye molecules. The detection of single molecules/chromophores or rhodamine-6G and equivalent dye molecules is suggested by system improvements. Thus, a com-

bination of improved light collection efficiency, improved detector quantum efficiency, or pulsed excitation and gated detection to reduce background noise can be used with the Nguyen et al. system. Detection of phycoerythrin was accomplished in the 180 μ s it took the molecule to flow through the focused laser beam.

In a preferred embodiment of the present process, the hydrodynamically focused flow system of Nguyen et al. is provided with an improved fluorescence detection system described in U.S Patent 4,793,705, Single Molecule Tracking, issued December 27, 1988, to Shera. As therein described, flow stream 24 provides to flow cell 26 modified nucleotides 14, 16, 18, and 22 in the sequence they are cleaved from DNA strand 10. Laser system 32 excites fluorescent dyes 14c, 16c, 18c, and 22c at selected wavelengths for identification in laminar sample flow 28 within flow cell 26.

Fluorescent events contained in optical signal 36 are focused by lens 38 on position sensitive detector system 42. Detector system 42 may comprise a microchannel plate (MCP) sensor to output spatial coordinates of observed photon events. An internal clock provides a temporal coordinate, wherein data processor 44 determines the presence of a molecule within flow cell 26. Molecular spectral response to laser 32 excitation enables the specific modified nucleotide to be identified. As noted by Shera, supra, data handling in the single molecule detection system 34 effectively provides a moving sample volume within focused flow stream 28 which contains only a single tagged nucleotide. System 34 can thus track multiple molecules existing within focused flow stream 28 to enable a high rate of sequencing to be maintained.

Referring now to Figure 2, there is shown a representative output signal from the single molecule detection system. The individual nucleotide molecules 14, 16, 18, and 22 are individually cleaved from DNA strand 10 into flow stream 24. The flow velocity and laminar flow conditions maintain the molecules in a train for sequential passage through flow cell 26 and the emitted photons from laser-excited molecular fluorescence are assigned to individual molecules passing within the cell. The characteristic dye for each type nucleotide is selected to have an identifiable excitation or fluorescence spectrum. This characteristic spectrum can be used to establish the base sequence for the DNA strand being investigated.

It will be appreciated that the present process further provides a capability to sort the detected molecules and deposit them on a moving substrate for subsequent identification, e.g., as described in M. R. Melamed et al., "Flow Cytometry and Sorting," Wiley, New York (1979). The flow stream maintains the bases spatially isolated in a flow

stream for presentation to a secondary identification device. The position between molecules on the moving substrate can be adjustable and can be large enough to resolve the sorted molecules by other techniques.

The foregoing description of the preferred embodiments of the invention has been presented for purposes of illustration and description.

10 **Claims**

1. A method for DNA and RNA base sequencing, comprising the steps of:
 - 15 hybridizing a single fragment of DNA or RNA to a substrate having a single site effective to hybridize with said fragment of DNA or RNA;
 - 20 introducing said single fragment into a flowing sample stream;
 - 25 sequentially cleaving the end base from the DNA or RNA fragment with exonuclease to form a train of said bases in the sample stream; and
 - 30 detecting said bases in said train in sequential passage through a detector which detects single molecules.
2. A method according to Claim 1, wherein each said base of said single fragment is modified to contain a tag having an identifiable characteristic for said base.
3. A method according to Claim 2, where said bases are modified prior to said cleavage.
- 35 4. A method according to Claim 2, further including the step of enzymatically synthesizing a strand of DNA complementary to a DNA or RNA strand to be characterized, where each nucleotide forming said synthesized strand contains a tag characteristic of that nucleotide.
5. A method according to Claim 2, wherein said tag is separated from the nucleotide by a linker arm effective for said cleavage.
- 45 6. A method according to Claim 1, wherein said cleaved bases are detected optically.
7. A method according to Claim 6, wherein each said tag is a fluorescent dye characteristic of one type of said nucleotides.
- 50 8. A method according to Claim 7, further including the step of exciting each said fluorescent dye and detecting the fluorescence spectrum of said dye.

9. A method according to Claim 1, further including the step of selecting said DNA or RNA fragments from a heterogeneous collection of DNA or RNA fragments wherein said site is a biotinylated probe effective to hybridize with DNA or RNA fragments to be selected.

10. A method for base sequencing of DNA or RNA fragments, comprising the steps of:

forming said fragments with each base having a characteristic identifiable fluorescent dye;

hybridizing a single fragment of DNA or RNA to a substrate having a single site effective to hybridize with said fragment of DNA or RNA;

sequentially cleaving single identifiable bases with exonuclease from the end of a single one of said fragments to form a train of said identifiable bases in a flowing sample stream; and

identifying said single cleaved bases in said train.

11. A method according to Claim 10, wherein the steps of forming said fragments includes the steps of forming by enzymatic synthesis a complementary strand of said DNA or RNA to be sequenced from said bases having identifiable fluorescent characteristics and thereafter base sequencing said complementary strand.

12. A method according to Claim 10, wherein said step of identifying said single, cleaved bases includes the step of exciting each said fluorescent dye and detecting the fluorescence spectrum of said dye.

13. A method for DNA or RNA base sequencing, comprising the steps of:

modifying each nucleotide effective for DNA or RNA synthesis to attach a fluorescent dye characteristic of that nucleotide with a linker arm effective to enable DNA or RNA synthesis and exonuclease cleavage;

hybridizing a single fragment of DNA or RNA to a substrate having a single site effective to hybridize with said fragment of DNA or RNA;

synthesizing from said modified nucleotides a strand of DNA complementary to a DNA or RNA strand having a base sequence to be determined;

cleaving each said modified nucleotide sequentially from the end of a single fragment containing said complementary DNA strand with exonuclease to form a train of said modified nucleotides in a flowing sample stream;

and

fluorescing each said characteristic dye to identify said sequence of nucleotides.

14. A method according to Claim 13, wherein the step of fluorescing said dyes further comprises the steps of:

exciting each said modified nucleotide with a laser effective to fluoresce said characteristic dye; and

detecting said fluorescence to sequentially identify said nucleotides and generate said sequence of said DNA or RNA.

15. A method according to Claim 13, further including the step of suspending a single fragment of synthesized DNA or RNA strand in a laminar flow stream.

16. A method according to Claim 13, wherein each synthesized DNA or RNA fragment is manipulated by hybridizing said fragment to a microsphere having a site effective for hybridization.

25 Patentansprüche

1. Verfahren zum DNS und RNS Basen-Sequenzieren, das folgenden Schritte aufweist:
Hybridisieren eines einzelnen DNS oder RNS Fragment zu einem Substrat, welches eine einzige Stelle hat, die geeignet ist, um mit dem DNS oder RNS Fragment zu hybridisieren;
Einführen dieses einzelnen Fragments in einen fließenden Probstrom;
sequentielles Abspalten der Endbase von dem DNS oder RNS Fragment mit einer Exonuklease, um eine Folge von diesen Basen in den Probstrom zu bilden; und
Erfassen oder Detektieren dieser Basen in dieser Folge in einem sequentiellen Durchfluß durch einen Detektor, welcher einzelne Moleküle erfaßt.
2. Verfahren nach Anspruch 1, wobei jede dieser Basen dieses einzelnen Fragments mit einer Kennung modifiziert ist, um ein identifizierbares Merkmal für jede Base zu haben.
3. Verfahren nach Anspruch 2, wobei die Basen vor der Abspaltung modifiziert werden.
4. Verfahren nach Anspruch 2, ferner aufweisend den Schritt des enzymatischen Synthetisierens eines DNS-Strang komplementär zu dem zu charakterisierenden DNS oder RNS Strang, wobei jedes Nukleotid, was den synthetisierten Strang bildet, eine Kennung aufweist, die für das Nukleotid charakteristisch ist.

5. Verfahren nach Anspruch 2, wobei jede Kennung getrennt ist von dem Nukleotid mittels eines Linker- bzw. Verbindungsarms geeignet für die Abspaltung.

6. Verfahren nach Anspruch 1, wobei die abgespaltenen Basen optisch erfaßt werden.

7. Verfahren nach Anspruch 6, wobei jede Kennung ein für jede Art dieser Nukleotide charakteristischer fluoreszenter Farbstoff ist.

8. Verfahren nach Anspruch 7, ferner aufweisend den Schritt des Anregens eines jeden der genannten fluoreszenten Farbstoffe und Erfassen des Fluoreszenzspektrums dieses Farbstoffs.

9. Verfahren nach Anspruch 1, ferner aufweisend den Schritt des Auswählens der genannten DNS oder RNS Fragmente aus einer heterogenen Sammlung von DNS oder RNS Fragmenten, wobei die genannte Stelle eine biotinversehene Probe ist, die geeignet ist, um mit den auszuwählenden DNS oder RNS Fragmenten zu hybridisieren.

10. Verfahren zum DNS und RNS Basen-Sequenzieren, das folgenden Schritte aufweist:
Bilden der Fragmente, wobei jede einzelne Base einen charakteristischen identifizierbaren fluoreszenten Farbstoff aufweist;
Hybridisieren eines einzelnen DNS oder RNS Fragments zu einem Substrat, welches eine einzelne Stelle hat, die geeignet ist, mit dem genannten DNS oder RNS Fragment zu hybridisieren;
sequentielles Abspalten einzelner identifizierbarer Basen mit einer Exonuclease vom Ende eines einzelnen Fragment, um einen Fluß von identifizierbaren Basen in einem fließenden Probestrom zu bilden; und
Identifizierung der einzelnen abgespaltenen Basen in dem Fluß.

11. Verfahren nach Anspruch 10, wobei die Schritte des Bildens dieser Fragmente die Schritte des Bildens durch enzymatische Synthese eines komplementären Strangs der DNS oder RNS einschließt, der von den Basen mit den identifizierbaren fluoreszenten Kennungen sequenziert wird und danach Basensequenzieren des komplementären Strangs.

12. Verfahren nach Anspruch 10, wobei der Schritt des Identifizierens der einzelnen, abgespaltenen Basen den Schritt des Anregens jedes der fluoreszenten Farbstoffe und das Erfassen des Fluoreszenzspektrums dieses Farbstoffs

5 beinhaltet.

13. Verfahren zum DNS und RNS Basen-Sequenzieren, das folgenden Schritte aufweist:
Modifizieren jedes Nukleotides geeignet für DNS oder RNS Synthese, in dem ein für jedes Nukleotide charakteristischer fluoreszenter Farbstoff mit einem Linkerarm verbunden wird, um DNS oder RNS Synthese und Exonuclease-Spaltung zu ermöglichen;
Hybridisieren eines einzelnen DNS oder RNS Fragments zu einem Substrat, welches eine einzelne Stelle geeignet zum Hybridisieren mit dem DNS oder RNS Fragment hat; Synthesieren eines zu dem DNS oder RNS Strang komplementären DNS Strang mit den modifizierten Nukleotiden, welcher die zu bestimmende Basensequenz aufweist;
Abspalten einzelner modifizierter Nukleotide sequentiell vom Ende des einzelnen Fragments, welches den komplementären DNS Strang enthält, mit einer Exonuclease, um einen Fluß von den modifizierten Nukleotiden in dem fließenden Probestrom zu bilden; und
Fluoreszieren jedes charakteristischen Farbstoffes, um die Sequenz der Nukleotide zu bestimmen oder identifizieren.

20

25

30

35

40

45

50

55

14. Verfahren nach Anspruch 13, wobei der Schritt des Fluoreszierens der Farbstoffe folgende Schritte aufweist: Anregen jedes der modifizierten Nukleotide mit einem Laser, der geeignet ist die charakteristischen Farbstoffe zu fluoreszieren, und
Erfassen dieser Fluoreszenz, um die Nukleotide sequentiell zu bestimmen und die Sequenz dieser DNS oder RNS zu erzeugen.

15. Verfahren nach Anspruch 13, ferner bestehend aus dem Schritt des Suspendierens eines einzelnen Fragments der synthetisierten DNS oder RNS in einen laminaren Stromfluß.

16. Verfahren nach Anspruch 13, wobei jedes synthetisierte DNS oder RNS Fragment manipuliert wird durch Hybridisieren dieses Fragments auf eine Microsphäre mit einer Stelle geeignet für die Hybridisierung.

Revendications

1. Procédé de séquençage de bases d'ADN et d'ARN, comprenant les étapes consistant :
à hybrider un seul fragment d'ADN ou d'ARN avec un substrat comportant un seul site efficace pour s'hybrider avec ce fragment d'ADN ou d'ARN;
à introduire cet unique fragment dans un

courant d'échantillon en circulation ;

à cliver successivement la base d'extrémité à partir du fragment d'ADN ou d'ARN, à l'aide d'une exonucléase pour former un train de ces bases dans le courant d'échantillon ; et

à détecter ces bases dans ce train par passage séquentiel à travers un détecteur qui détecte de simples molécules.

2. Procédé selon la revendication 1, dans lequel chaque base de ce simple fragment est modifiée de façon à comporter un marqueur ayant une caractéristique identifiable pour cette base. 10

3. Procédé selon la revendication 2, dans lequel les bases sont modifiées avant le clivage. 15

4. Procédé selon la revendication 2, comprenant en Outre l'étape de synthèse enzymatique d'un brin d'ADN complémentaire d'un brin d'ADN ou d'ARN destiné à être caractérisé, chaque nucléotide formant le brin synthétisé comportant un marqueur caractéristique de ce nucléotide. 20

5. Procédé selon la revendication 2, dans lequel le marqueur est séparé du nucléotide par un bras de liaison efficace pour le clivage. 25

6. Procédé selon la revendication 1, dans lequel les bases clivées sont détectées par voie optique. 30

7. Procédé selon la revendication 6, dans lequel chaque marqueur est un colorant fluorescent caractéristique d'un type de ces nucléotides. 35

8. Procédé selon la revendication 7, comprenant en outre l'opération d'excitation de chacun de ces colorants fluorescents et de détection du spectre de fluorescence du colorant. 40

9. Procédé selon la revendication 1, comprenant en outre l'étape de sélection des fragments d'ADN ou d'ARN à partir d'une collection hétérogène de fragments d'ADN ou d'ARN, le site consistant en une sonde biotinylée efficace pour s'hybrider avec les fragments d'ADN ou d'ARN destinés à être sélectionnés. 45

10. Procédé de séquençage de bases de fragments d'ADN ou d'ARN, comprenant les étapes consistant :

à former ces fragments, chaque base comportant un colorant fluorescent caractéristique identifiable ;

à hybrider un seul fragment d'ADN ou d'ARN avec un substrat comportant un seul site efficace pour s'hybrider avec ce fragment d'ADN ou d'ARN ;

à cliver successivement de simples bases identifiables à l'aide d'une exonucléase à partir de l'extrémité d'un seul de ces fragments, afin de former un train desdites bases identifiables dans un courant d'échantillon en circulation ; et

à identifier ces simples bases clivées dans ledit train.

11. Procédé selon la revendication 10, dans lequel les étapes de formation des fragments comprennent la formation par synthèse enzymatique d'un brin complémentaire de l'ADN ou de l'ARN destiné à être séquencé à partir des bases ayant des caractéristiques de fluorescence identifiables, et le séquençage ultérieur des bases de ce brin complémentaire. 5

12. Procédé selon la revendication 10, dans lequel l'étape d'identification des simples bases clivées comprend l'opération d'excitation de chacun des colorants fluorescents, et de détection du spectre de fluorescence de ce colorant. 10

13. Procédé de séquençage de bases d'ADN ou d'ARN, comprenant les étapes consistant :

à modifier chaque nucléotide efficace pour la synthèse d'ADN ou d'ARN, par liaison d'un colorant fluorescent caractéristique de ce nucléotide, à l'aide d'un bras de liaison efficace pour permettre la synthèse d'ADN ou d'ARN et le clivage par une exonucléase ;

à hybrider un seul fragment d'ADN ou d'ARN avec un substrat comportant un seul site efficace pour s'hybrider avec ce fragment d'ADN ou d'ARN ;

à synthétiser à partir de ces nucléotides modifiés un brin d'ADN complémentaire d'un brin d'ADN ou d'ARN ayant une séquence de bases destinée à être déterminée ;

à cliver chacun de ces nucléotides modifiés successivement à partir de l'extrémité d'un seul fragment contenant le brin d'ADN complémentaire, à l'aide d'une exonucléase afin de former un train de ces nucléotides modifiés dans un courant d'échantillon en circulation ; et

à rendre fluorescent chacun de ces colorants caractéristiques afin d'identifier la séquence de nucléotides. 15

14. Procédé selon la revendication 13, dans lequel l'étape consistant à rendre fluorescents ces colorants comprend en outre :

l'excitation de chacun des nucléotides modifiés avec un laser efficace pour rendre fluorescent le colorant caractéristique ; et

la détection de la fluorescence pour identifier successivement lesdits nucléotides et générer la séquence de l'ADN ou de l'ARN.

15. Procédé selon la revendication 13, comprenant en outre l'opération de mise en suspension d'un seul fragment de brin d'ADN ou d'ARN synthétisé, dans un courant laminaire. 5
16. Procédé selon la revendication 13, dans lequel chaque fragment d'ADN ou d'ARN synthétisé est manipulé par hybridation de ce fragment avec une microsphère comportant un site efficace pour une hybridation. 10

15

20

25

30

35

40

45

50

55

11

Fig. 1

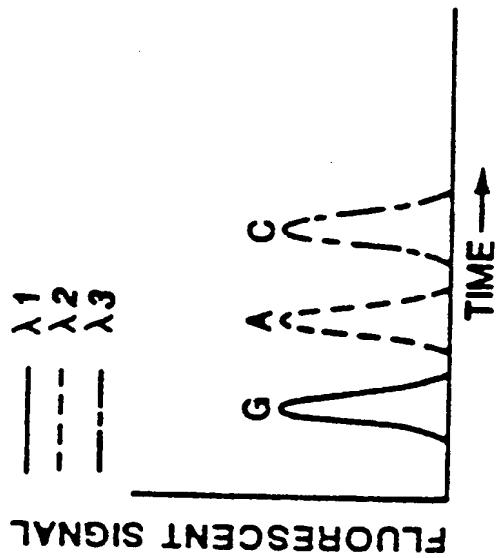
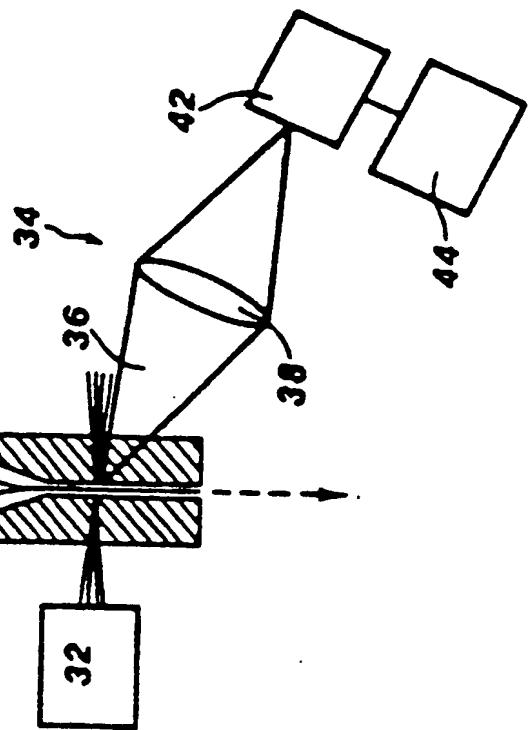


Fig. 2

